



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 08:16 pm BST

PDB ID : 5CER
Title : Bd0816 Predatory Endopeptidase from *Bdellovibrio bacteriovorus* in complex with immunity protein Bd3460
Authors : Lovering, A.L.; Cadby, I.T.; Lambert, C.; Sockett, R.E.
Deposited on : 2015-07-07
Resolution : 2.48 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

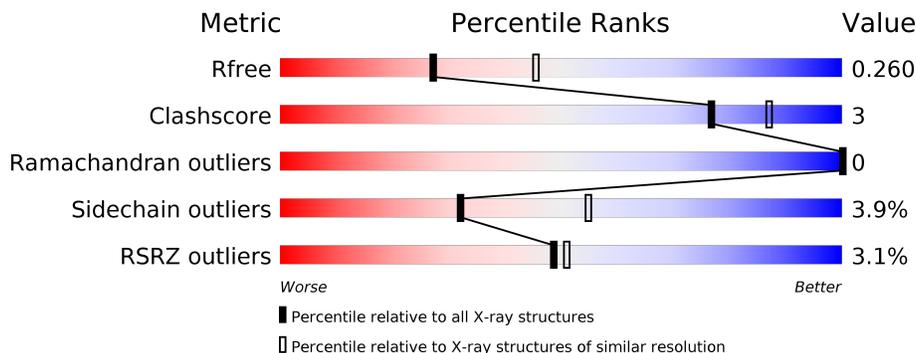
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.48 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



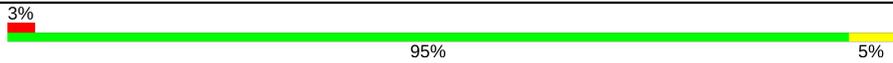
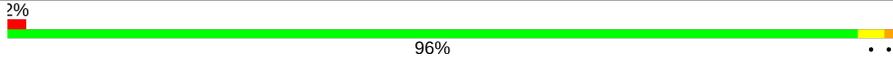
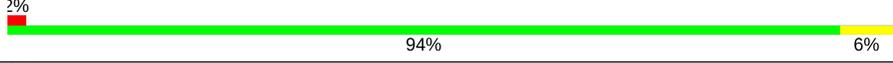
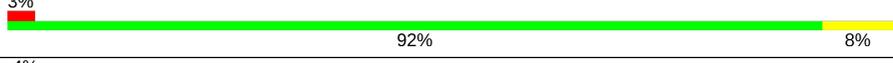
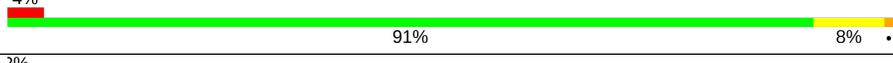
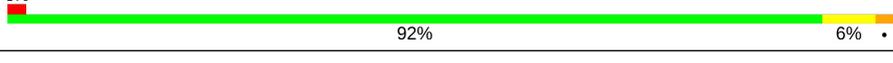
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	401	 2% 90% 9%
1	C	401	 5% 90% 7%
1	E	401	 % 88% 11%
1	G	401	 6% 89% 10%
1	I	401	 3% 90% 9%
1	K	401	 2% 88% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	193	 3% 95% 5%
2	D	193	 2% 96% ..
2	F	193	 2% 94% 6%
2	H	193	 3% 92% 8% .
2	J	193	 4% 91% 8% ..
2	L	193	 2% 92% 6% .

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 28724 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Bd0816.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	401	3124	1970	536	604	14	0	0	0
1	C	401	3124	1970	536	604	14	0	0	0
1	E	401	3124	1970	536	604	14	0	0	0
1	G	399	3110	1962	534	600	14	0	0	0
1	I	401	3124	1970	536	604	14	0	0	0
1	K	400	3124	1971	535	603	15	0	1	0

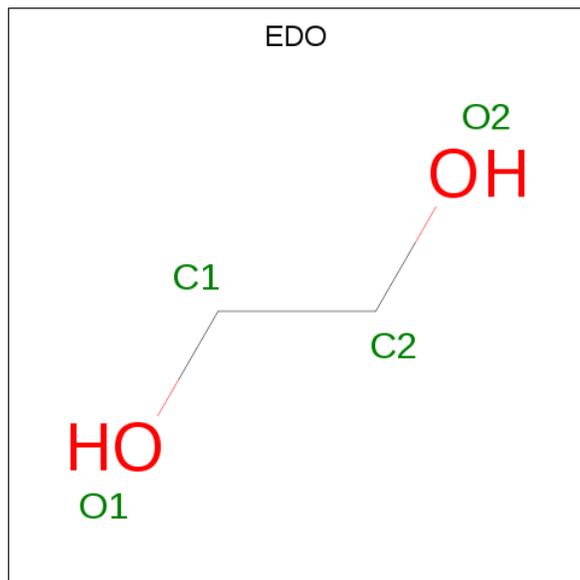
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	MET	LYS	engineered mutation	UNP Q6MPN2
A	58	ALA	SER	engineered mutation	UNP Q6MPN2
C	26	MET	LYS	engineered mutation	UNP Q6MPN2
C	58	ALA	SER	engineered mutation	UNP Q6MPN2
E	26	MET	LYS	engineered mutation	UNP Q6MPN2
E	58	ALA	SER	engineered mutation	UNP Q6MPN2
G	26	MET	LYS	engineered mutation	UNP Q6MPN2
G	58	ALA	SER	engineered mutation	UNP Q6MPN2
I	26	MET	LYS	engineered mutation	UNP Q6MPN2
I	58	ALA	SER	engineered mutation	UNP Q6MPN2
K	26	MET	LYS	engineered mutation	UNP Q6MPN2
K	58	ALA	SER	engineered mutation	UNP Q6MPN2

- Molecule 2 is a protein called Bd3460.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			
2	D	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			
2	F	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			
2	H	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			
2	J	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			
2	L	193	Total	C	N	O	S	0	0	0
			1439	897	250	288	4			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	I	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	181	Total	O	0	0
			181	181		
4	B	93	Total	O	0	0
			93	93		
4	C	141	Total	O	0	0
			141	141		

Continued on next page...

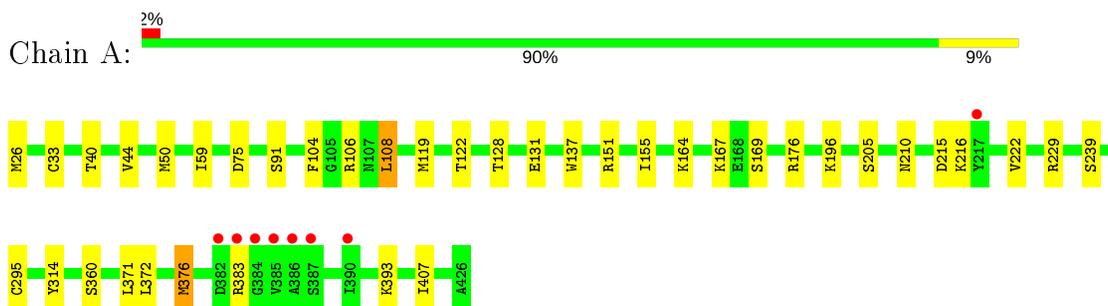
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	94	Total O 94 94	0	0
4	E	141	Total O 141 141	0	0
4	F	77	Total O 77 77	0	0
4	G	112	Total O 112 112	0	0
4	H	84	Total O 84 84	0	0
4	I	131	Total O 131 131	0	0
4	J	69	Total O 69 69	0	0
4	K	136	Total O 136 136	0	0
4	L	97	Total O 97 97	0	0

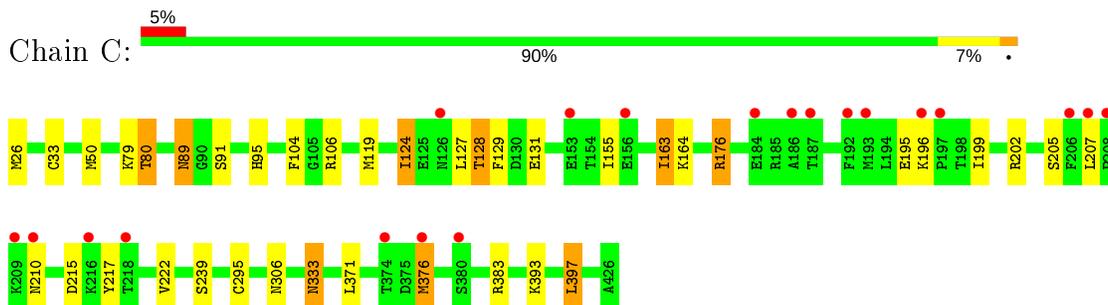
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

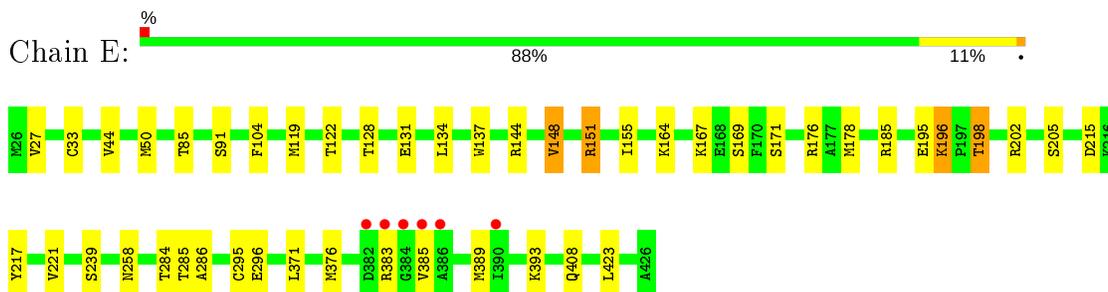
- Molecule 1: Bd0816



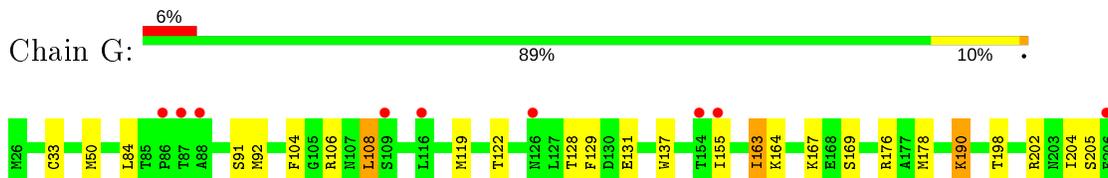
- Molecule 1: Bd0816



- Molecule 1: Bd0816

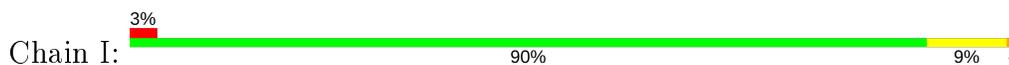


- Molecule 1: Bd0816

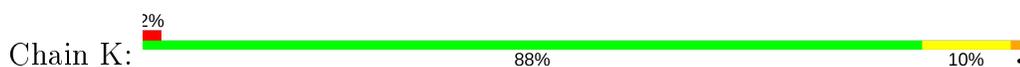




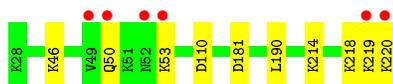
- Molecule 1: Bd0816



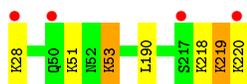
- Molecule 1: Bd0816



- Molecule 2: Bd3460



- Molecule 2: Bd3460



- Molecule 2: Bd3460

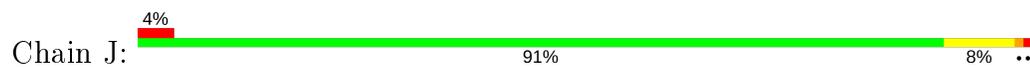


- Molecule 2: Bd3460

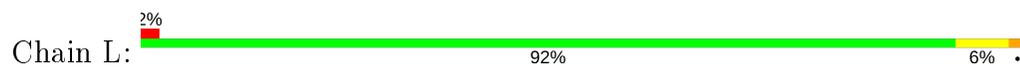




- Molecule 2: Bd3460



- Molecule 2: Bd3460



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	237.45Å 212.32Å 90.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	100.00 – 2.48 103.62 – 2.48	Depositor EDS
% Data completeness (in resolution range)	98.8 (100.00-2.48) 98.8 (103.62-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.16	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.73 (at 2.48Å)	Xtrriage
Refinement program	REFMAC 5.8.0103	Depositor
R, R_{free}	0.215 , 0.249 0.233 , 0.260	Depositor DCC
R_{free} test set	7978 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	31.0	Xtrriage
Anisotropy	0.481	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 43.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	28724	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.02% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality i

5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: EDO

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.71	0/3176	0.84	5/4287 (0.1%)
1	C	0.72	1/3176 (0.0%)	0.88	8/4287 (0.2%)
1	E	0.73	1/3176 (0.0%)	0.89	12/4287 (0.3%)
1	G	0.64	0/3162	0.86	10/4268 (0.2%)
1	I	0.73	1/3176 (0.0%)	0.86	7/4287 (0.2%)
1	K	0.69	0/3179	0.89	10/4290 (0.2%)
2	B	0.69	1/1447 (0.1%)	0.77	2/1940 (0.1%)
2	D	0.63	0/1447	0.79	1/1940 (0.1%)
2	F	0.61	0/1447	0.78	2/1940 (0.1%)
2	H	0.63	0/1447	0.78	3/1940 (0.2%)
2	J	0.63	0/1447	0.83	4/1940 (0.2%)
2	L	0.65	0/1447	0.79	0/1940
All	All	0.68	4/27727 (0.0%)	0.85	64/37346 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
1	E	0	1
1	K	0	1
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	171	SER	CA-CB	7.87	1.64	1.52
1	I	91	SER	CB-OG	-6.99	1.33	1.42
1	C	333	ASN	C-N	6.23	1.48	1.34

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	53	LYS	CD-CE	5.11	1.64	1.51

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	K	223	LEU	CB-CG-CD2	-11.33	91.74	111.00
1	E	185	ARG	CG-CD-NE	9.56	131.89	111.80
1	C	124	ILE	CG1-CB-CG2	-9.48	90.53	111.40
1	I	383	ARG	CG-CD-NE	8.72	130.11	111.80
1	E	178	MET	CG-SD-CE	8.59	113.94	100.20

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	333	ASN	Mainchain
1	E	196	LYS	Peptide
1	K	424	THR	Peptide

5.2 Too-close contacts [\(i\)](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3124	0	3148	18	0
1	C	3124	0	3148	24	0
1	E	3124	0	3148	25	1
1	G	3110	0	3137	23	0
1	I	3124	0	3148	25	1
1	K	3124	0	3152	29	0
2	B	1439	0	1499	3	0
2	D	1439	0	1499	3	1
2	F	1439	0	1499	3	0
2	H	1439	0	1499	8	0
2	J	1439	0	1499	7	1
2	L	1439	0	1499	7	0
3	I	4	0	6	3	0
4	A	181	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	93	0	0	0	0
4	C	141	0	0	6	0
4	D	94	0	0	1	0
4	E	141	0	0	1	0
4	F	77	0	0	0	0
4	G	112	0	0	0	0
4	H	84	0	0	1	0
4	I	131	0	0	1	0
4	J	69	0	0	0	0
4	K	136	0	0	1	0
4	L	97	0	0	0	0
All	All	28724	0	27881	168	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

The worst 5 of 168 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:385:VAL:HG12	1:E:389:MET:CE	1.78	1.13
1:I:376:MET:HE3	1:I:383:ARG:NH1	1.83	0.94
1:A:75:ASP:OD1	1:A:229:ARG:NH1	2.08	0.85
1:E:385:VAL:HG12	1:E:389:MET:HE2	1.57	0.85
1:I:376:MET:CE	1:I:383:ARG:NH1	2.46	0.78

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:220:LYS:NZ	2:J:162:THR:OG1[2_455]	2.08	0.12
1:E:151:ARG:NH2	1:I:202:ARG:NE[4_554]	2.17	0.03

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	399/401 (100%)	390 (98%)	9 (2%)	0	100	100
1	C	399/401 (100%)	388 (97%)	11 (3%)	0	100	100
1	E	399/401 (100%)	389 (98%)	10 (2%)	0	100	100
1	G	397/401 (99%)	389 (98%)	8 (2%)	0	100	100
1	I	399/401 (100%)	390 (98%)	9 (2%)	0	100	100
1	K	399/401 (100%)	390 (98%)	9 (2%)	0	100	100
2	B	191/193 (99%)	187 (98%)	4 (2%)	0	100	100
2	D	191/193 (99%)	189 (99%)	2 (1%)	0	100	100
2	F	191/193 (99%)	189 (99%)	2 (1%)	0	100	100
2	H	191/193 (99%)	188 (98%)	3 (2%)	0	100	100
2	J	191/193 (99%)	187 (98%)	4 (2%)	0	100	100
2	L	191/193 (99%)	187 (98%)	4 (2%)	0	100	100
All	All	3538/3564 (99%)	3463 (98%)	75 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	340/340 (100%)	325 (96%)	15 (4%)	28	49
1	C	340/340 (100%)	325 (96%)	15 (4%)	28	49
1	E	340/340 (100%)	327 (96%)	13 (4%)	33	56
1	G	339/340 (100%)	325 (96%)	14 (4%)	30	53
1	I	340/340 (100%)	322 (95%)	18 (5%)	22	40
1	K	341/340 (100%)	324 (95%)	17 (5%)	24	43
2	B	150/150 (100%)	149 (99%)	1 (1%)	84	93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	150/150 (100%)	146 (97%)	4 (3%)	44	69
2	F	150/150 (100%)	146 (97%)	4 (3%)	44	69
2	H	150/150 (100%)	148 (99%)	2 (1%)	69	86
2	J	150/150 (100%)	143 (95%)	7 (5%)	26	46
2	L	150/150 (100%)	145 (97%)	5 (3%)	38	61
All	All	2940/2940 (100%)	2825 (96%)	115 (4%)	32	55

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	91	SER
1	G	371	LEU
1	K	371	LEU
1	G	104	PHE
1	G	190	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	57	ASN
2	H	158	ASN
1	K	373	HIS
2	H	59	GLN
2	H	75	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	EDO	I	501	-	3,3,3	1.15	0	2,2,2	0.96	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	I	501	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	501	EDO	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	401/401 (100%)	-0.16	8 (1%) 65 67	17, 31, 66, 143	0
1	C	401/401 (100%)	0.14	20 (4%) 28 30	15, 36, 99, 146	0
1	E	401/401 (100%)	-0.00	6 (1%) 73 75	17, 37, 74, 128	0
1	G	399/401 (99%)	0.21	23 (5%) 23 23	17, 42, 95, 224	0
1	I	401/401 (100%)	-0.03	14 (3%) 44 46	10, 34, 79, 181	0
1	K	400/401 (99%)	-0.02	10 (2%) 57 59	14, 35, 85, 129	0
2	B	193/193 (100%)	-0.03	6 (3%) 49 51	24, 38, 72, 141	0
2	D	193/193 (100%)	-0.20	4 (2%) 63 65	24, 37, 68, 100	0
2	F	193/193 (100%)	-0.08	3 (1%) 72 73	28, 43, 83, 120	0
2	H	193/193 (100%)	-0.03	5 (2%) 56 58	27, 40, 68, 111	0
2	J	193/193 (100%)	0.07	7 (3%) 42 45	26, 47, 82, 118	0
2	L	193/193 (100%)	-0.18	3 (1%) 72 73	23, 40, 82, 131	0
All	All	3561/3564 (99%)	-0.01	109 (3%) 49 51	10, 38, 84, 224	0

The worst 5 of 109 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	220	LYS	15.4
2	F	220	LYS	7.3
1	G	210	ASN	5.7
2	B	52	ASN	5.6
2	J	220	LYS	5.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
3	EDO	I	501	4/4	0.79	0.25	31,37,37,40	0

6.5 Other polymers [i](#)

There are no such residues in this entry.